

703017

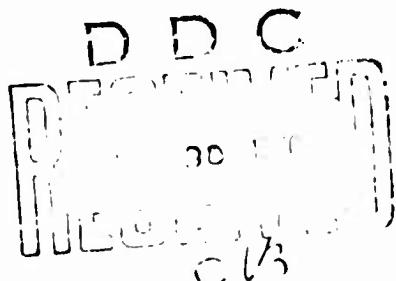
RESEARCH STUDY

EVALUATION & ANALYSIS BRANCH

AN INVESTIGATION OF PEER-RATINGS vs
SELF-RATINGS vs SUPERVISORY RATINGS

NUMBER 33

DATE November 1965



U S ARMY

ENLISTED EVALUATION CENTER

FORT BENJAMIN HARRISON
INDIANAPOLIS, INDIANA

Research Study No. 33

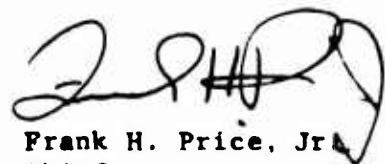
**AN INVESTIGATION OF PEER-RATINGS vs
SELF-RATINGS vs SUPERVISORY RATINGS**

Ronald K. Goodnight

Submitted by

**Dale R. Baker
Chief, Analysis Section**

Approved by



**Frank H. Price, Jr.
Chief
Evaluation and Analysis Branch**


**James C. Donaghey
Colonel, Infantry
Commanding**

November 1965

BLANK PAGE

Table of Contents

| | |
|--|----------|
| Introduction. | 1 |
| Method. | 2 |
| Results | 3 |
| Table I - Intercorrelation Matrix Between Three Rating Methods on Overall Performance | 4 |
| Table II - Intercorrelation Matrix Between Three Rating Methods on Promotion Qualification | 4 |
| Table III - Correlations Between Overall Performance Scale and Promotion Qualification Scale from Three Rating Methods. | 4 |
| Discussion. | 5 |
| Conclusion. | 5 |
| References. | 6 |
| Appendix | 7 |

Introduction

One of the foremost problems in psychological measurement of personnel performance is the selection of a reliable, as well as valid rating method. An equally large problem, once the method is selected, is who should accomplish the evaluations.

The US Army Enlisted Evaluation Center utilizes peer or co-worker ratings as a criterion for validation of Evaluation Tests (ET) and Commander's Evaluation Reports (CER). This particular rating method was selected in lieu of other methods because: 1) numerous journal articles and books (Hollander, 1954a, 1954b, 1956a, 1956b, 1957; Doll & Longo, 1962; Turk, 1961; Tiffin & McCormick, 1960; Cornbach, 1960; Whisler & Harper, 1962) advocate and support the superiority of peer-ratings, particularly their reliability and validity; 2) supervisory criterion ratings would overlap with the supervisor's CER ratings; and 3) the use of peer-ratings at USAEEC have proven to be the most reliable as well as practical criterion.

In many instances, however, it is not possible to obtain sufficient peer-ratings to be used as a criterion. This is particularly true at the higher Military Occupational Speciality (MOS) skill levels in which an insufficient number of higher grade personnel work closely enough together to be qualified to rate one another. A possible solution to this problem is to use either self-ratings for validation of the ET and CER or supervisory ratings as a criterion solely for the validation of the ET.

Purpose. Thus, it is the objective of this study to compare peer-ratings, self-ratings, and supervisory ratings to determine if the self or supervisory ratings correlate significantly to the peer-ratings to warrant their use as a criterion.

Method

Peer, self, and supervisory ratings were obtained during the period 12 - 16 April 1965 on 215 enlisted men stationed at Fort Riley, Kansas. Samples were selected from 36 separate MOS skill levels.

Three peer-ratings were rendered on each ratee by enlisted men within the same MOS skill level. Each supervisor accomplishing the ratings was the immediate superior of the ratee. All raters had observed the man's on-the-job performance at least several times a week and had worked with or supervised the man a minimum of one month. The purpose of the ratings was explained to the raters and the need for rendering accurate ratings was emphasized.

The ratings were accomplished on two factors, "Overall Performance" and "Promotion Qualification," using a discreet graphic scale ranging from 1 (poor) to 11 (exceptional) as shown on EPEECE Form 70 (see Appendix).

Pearson product-moment correlation coefficients were computed between the three types of ratings on both factors. These correlations were then tested for significant differences. The correlations were also computed between the two rating factors using each of the rating types, respectively, to determine the amount of agreement between the scales.

Results

An intercorrelation matrix for the three types of ratings obtained on the "Overall Performance" scale is shown in Table I. The intercorrelation matrix for the "Promotion Qualification" scale is presented in Table II. The highest correlation coefficients were between the peer-ratings and the supervisors' ratings (.3922 and .4766) on the two scales, respectively. Whereas, the lowest correlations on both rating scales were between the self-ratings and the supervisory ratings (.2463 and .2397). The reliability of the peer-ratings was .4625, significant at the .001 level.

The correlation coefficients between the two scales computed from each type of rating are shown in Table III. Each correlation coefficient was significantly different from zero at the .001 significance level. The two scales were most highly correlated (.8494) using the self-ratings and the least correlation (.7296) was found using the peer-ratings.

Table I

Intercorrelation Matrix Between Three Rating Methods on Overall Performance*

| | | Rating Method | | |
|---------------|------------|---------------|-------|-------------|
| | | Peer | Self | Supervisory |
| Rating Method | Peer | -- | .2536 | .3922 |
| | Self | | -- | .2463 |
| | Supervisor | | | -- |

Table II

Intercorrelation Matrix Between Three Rating Methods on Promotion Qualification*

| | | Rating Method | | |
|---------------|------------|---------------|-------|-------------|
| | | Peer | Self | Supervisory |
| Rating Method | Peer | -- | .3111 | .4766 |
| | Self | | -- | .2397 |
| | Supervisor | | | -- |

Table III

Correlations Between Overall Performance Scale and Promotion Qualification Scale from Three Rating Methods*

| | | Rating Scales |
|---------------|------------|---------------|
| | | OP vs PQ** |
| Rating Method | Peer | .7296 |
| | Self | .8494 |
| | Supervisor | .7731 |

*All correlation coefficients are significant at the .01 level.
**Overall Performance (OP) vs Promotion Qualification (PQ).

Discussion

The purpose of this study was to investigate and compare peer-ratings, the present criterion, with self-ratings, and supervisory ratings. If either the self or supervisory ratings were significantly correlated with the peer-ratings, they could be substituted as the criterion when an insufficient number of peer-ratings are available. If the self-ratings agreed significantly with the peer-ratings, the self-ratings could then be used as the criterion for validation of the ET and CER when peer-ratings were not obtainable in sufficient quantity. In addition, if the supervisory ratings were significantly correlated with the peer-ratings, then the supervisory ratings could be used as the criterion for validation of the ET.

The intercorrelations between the various rating methods on the "Overall Performance" scale as shown in Table I reveals that the supervisory ratings were the most highly related to the peer-ratings having a correlation coefficient of .3922. Although the correlation coefficients .2536 for self-rating vs peer-ratings and .2463 for self-rating vs supervisory rating, respectively were somewhat lower than the peer vs supervisor correlation, no significant differences were found between the correlation coefficients at the .05 significance level. However, all the correlation coefficients were significantly different from zero at the .01 level.

The intercorrelation coefficients between the rating methods on the "Promotion Qualification" scale are shown in Table II. Again the highest correlation (.4766) was between the peer-ratings and the supervisory ratings. The correlation coefficients were .3111 and .2397 for the peer vs self ratings and self vs supervisory ratings, respectively. All correlation coefficients were significantly different from zero at the .01 level. The peer vs supervisor correlation (.4766) was significantly different from the self vs supervisor correlation at the .05 level ($t = 2.4227$). No significant differences were found between the remaining correlations.

The correlation coefficients presented in Table III show that the least difference between the two rating scales existed using self-ratings with a correlation of .8489. The greatest difference ($r = .7296$) was found with the peer-ratings. Although no significant differences were found between these coefficients, it appears that possibly the peer raters make somewhat greater discriminations between the two rating factors than do the self or supervisory raters.

Conclusion

On the basis of this study it appears that supervisory ratings would be preferable to self-ratings as a substitute criterion when peer-ratings are inaccessible. However, further and more extensive research is needed in this area before such criterion substitution and implementation could be incorporated in a validation program.

References

Cronbach, L. J. Essentials of psychological testing. (2nd ed.).
New York: Harper & Brothers, 1960.

Doll, R. E. and Longo, A. A. Improving the predictive effectiveness
of peer ratings. Personnel psychol., 1962, 15(2), 215-220.

Hollander, E. P. Buddy ratings: military research and industrial
implications. Personnel psychol., 1954a 7, 385-393.

Hollander, E. P. Peer nominations on leadership as a predictor of the
pass-fail criterion in naval air training. J. appl. psychol.,
1954b, 38, 150-153.

Hollander, E. P. Interpersonal exposure time as a determinant of the
predictive utility of peer ratings. Psychol. rep., 1956a, 2,
445-448.

Hollander, E. P. The friendship factor in peer nominations. Personnel
psychol., 1956b, 9, 435-447.

Hollander, E. P. The reliability of peer nominations under various
conditions of administration. J. appl. psychol., 41, 85-90.

Tiffin, J., and McCormick, E. J. Industrial psychology. (4th ed.).
New Jersey: Prentice-Hall, 1960.

Turk, H. Instrumental and expressive ratings reconsidered. Sociometry,
1961, 24, 76-81.

Whisler, T. L., and Harper, S. F. Performance appraisal. New York:
Holt, Rinehart, & Winston, 1962.

INSTRUCTIONS TO RATER

The ratings you give will be used for research purposes only. No record will be kept of your ratings. Take them as accurate as you can, since they will help improve the Army Enlisted Evaluation System. Complete blanks A, B, and C first.

| | |
|---|--------------|
| A. Your duty assignment and station. | B. Your MOS. |
| C. Your name, grade, service number and the date. | |

BEFORE COMPLETING THE RATING FORM READ THESE DIRECTIONS CAREFULLY: At the bottom of this form are six columns for you to complete. Notice that the name, service number, PMOS and pay grade of the men you will rate has been written in columns D, E, and F. For each man answer questions 1, 2, 3, 4, 5, and 6 in the numbered columns with the appropriate letter or number. Refer to the questions numbered 1-6 below, but answer beside each name in the column corresponding to the number of the question.

SPECIAL DIRECTIONS FOR ITEMS 5 AND 6. For each question write in the column the one number which best describes the man you are rating. The NUMBER "1" SHOULD BE USED FOR THE MEN WHO ARE POOR IN THEIR MOS AS COMPARED WITH ALL THE OTHERS YOU HAVE KNOWN. THE NUMBER "11" SHOULD BE USED ONLY FOR THE MEN WHO ARE THE VERY BEST YOU HAVE KNOWN. Obviously, most men will be rated in the "Adequate," "Good," and "Superior" areas--numbers 3-9. But a man may be definitely exceptional and properly be rated number 10 or 11, or definitely poor and properly rated number 1 or 2. Also, a man may be rated low on one question and high on the other question. Rate each man separately on both questions and rate on each question independently. Do not hesitate to use the lower numbers when a man is poor or adequate, or the higher numbers when a man is superior or exceptional.

REMEMBER: The GOOD rater--Takes his TIME in making a rating, Rates SEPARATELY on each quality--not on general impression, Bases his judgment on USUAL performance, Rates EACH man separately from other men.

| | |
|--|---|
| 1. What is your basis for judging this man's performance? | 2. How long have you worked with or supervised him? |
| a. I see him at work all the time b. I see him at work several times a day c. I see him at work several times a week d. I seldom see him in the work situation | a. under one month b. 1 to 2 months c. 3 to 5 months d. 6 months or over |
| 3. What is your relationship to this man? | 4. To the best of your knowledge, is this man actually performing the normal duties of his Primary MOS? |
| a. fellow worker b. immediate supervisor (officer or noncommissioned officer directing and assigning man's work) c. higher level supervisor (supervisor of his immediate supervisor) | a. yes b. no |
| 5. In your opinion, how is this man's overall performance in his Primary Four-Digit MOS? | |

| | | | | | | | | | | |
|------|----------|------|----------|-------------|---|---|---|---|----|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| Poor | Adequate | Good | Superior | Exceptional | | | | | | |

3. What is your relationship to this man?

- a. fellow worker
- b. immediate supervisor (officer or non-officer directing and assigning work)
- c. higher level supervisor (supervisor)

4. To the best of your knowledge, is this man actually performing the normal duties of his Primary MOS?

a. yes
b. no

5. In your opinion, how is this man's overall performance in his Primary Four-Digit MOS?

6. In your opinion, how well would this man do in the next higher grade?

APPENDIX

**UNITED STATES ARMY
ENLISTED PERSONNEL RATING FORM**



**The U S Army Enlisted Evaluation Center
Fort Benjamin Harrison
Indianapolis, Indiana 46249**

**FOR OFFICIAL USE ONLY
(when completed)**